

調查資料管理與共享規劃檢核表

壹、簡介

一、目的或使用機會

「調查資料管理與共享規劃檢核表」，係由中央研究院調查研究專題中心建置的「學術調查研究資料庫」（以下簡稱本資料庫）編製。主要目的是為有意以調查方式蒐集資料的研究者，對於資料的管理、分享，能預先參照檢核表的內容，做好妥善的規劃。對於資料的內容與價值、資料品質的管理、應注意的倫理議題、資料分享的規劃等面向，此一檢核表不僅羅列其中涉及的問題，並提供研究者可行的解決方向與建議。

二、背景說明

由於社會科學研究方法的不斷創新及資訊科技的進步，運用調查資料來進行學術研究，有益趨風行的趨勢，也帶動國內調查計畫的數量明顯增加。並且，隨著社會科學、公衛醫療領域跨國及/或長期（追蹤）研究的發展趨勢，近年來國內這類調查計畫的數量也明顯增多。然而，另一方面，由於個人隱私權愈益受到重視，使得研究倫理、個人資料的保密問題，也對蒐集資料的研究者形成挑戰。

從成本、效益的角度而言，這些調查計畫，尤其是長期性、大規模的調查，所蒐集的調查資料內容豐富，蘊藏著多種可能的研究主題。如果這些資料只能被調查計畫團隊使用，會使得消耗大量人力、物力蒐集到的資料無法被充分應用。反之，如果資料蒐集團隊能與他人共享資料，對團隊本身有多項好處：1. 可促進與其他研究者間的合作，提昇研究成果的品質；2. 能提高研究成果能見度並具更大的影響力；3. 當其他使用資料的研究者引述資料來源時，可增加原研究團隊成員的學術聲譽。因此，將調查計畫所得到的資料，於合理的時間內開放學界共享，是許多先進國家的共識。例如，歐美一些國家的研究經費補助機關，相繼要求研究者於提出研究計畫與經費申請的同時，也提出資料管理計畫，並納入審查項目，以確保資料公開、資源共享，並保障受訪者的權益。

本資料庫的建置目的之一，即是為了推動調查資料的再次應用，為協助國內調查計畫預先做好資料管理與共享規劃，並倡導資料共享的理念，我們參酌了多

個國際組織的文件，進而考量國內研究環境，編製出「調查資料管理與共享規劃檢核表」，提供擬進行調查計畫的研究者參考使用。所參考的文件，包括英國 Digital Curation Centre (DCC) 研發的 DCC Data Management Planning Tool 中的“Data Management Plans”，以及 UK Data Archive (UKDA) 的“Managing and Sharing Data-Best Practice for Researchers”（2011 年 3 月版）等。

檢核表除列出研究者在規劃調查資料的管理與共享時應考量的面向外，也就各面向提出因應對策。整體而言，研究者如果能預先參照本檢核表進行規劃，有下列好處：

1. 可事先設計資料的管理方式，不但能提高資料正確性與研究品質，也便於未來資料與資訊的搜尋。
2. 可事先考量資料的隱私問題，避免蒐集不必要的隱私資料，或誤刪需要保留的資料。
3. 避免重複蒐集資料，或避免重複進行相同工作浪費人力。

三、檢核表基本架構

檢核表內容主要包含六大部分，分別為：

1. 計畫背景介紹
2. 倫理道德與隱私議題
3. 資料分享
4. 計畫執行期間的資料儲存及管理
5. 計畫結束後的資料儲藏、保存及管理
6. 管理需求與確認

貳、檢核表

「調查資料管理與共享計畫」內容檢核表

	編號	項目	說明
	1	計畫背景介紹	
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	1.1	計畫基本資訊說明	a. 是否為延續性計畫 b. 是否會持續追蹤訪問 c. 計畫其他參與人員的基本資料與聯絡方式 d. 將資料公開與他人共享的規劃或協議，例如國科會補助之調查計畫，需繳交一份資料由中央研究院人文社會科學研究中心調查研究專題中心之「學術調查研究資料庫」公開或保存。 e. 計畫目的
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	1.2	在資料分享、著作權或其他智慧財產權上有無正式的協議？	a. 誰擁有資料的著作權？著作權是否屬於共同所有？ b. 多方(研究人員或機構/單位)合作計畫的合約書中，應包括著作權及其他智慧財產權歸屬的協議，該協議內容亦須考量計畫執行期間機構/單位人員流動的可能性，以及爭議發生時的調解程序或機制。 c. 受訪者完訪資料的智慧財產權歸屬。
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	1.3	要進行計畫的研究問題，是否有現存資料可用？	先確定無現存資料可供使用，以免重複蒐集資料，耗時費力。
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	1.4	若有現存資料，但無法取得，說明面臨的問題	問題可能是管道或費用等。
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	1.5	為何需要蒐集新的資料？	可能的原因包括：沒有合適的現存資料、想將研究拓展至其他領域、新資料將用來與舊有資料做比較等。

<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	1.6	資料蒐集方法	可能是方式包括：面對面訪問、電話訪問、網路調查、自填問卷、郵寄問卷等等。
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	1.7	如何確保資料品質？	<p>可由多方面確保未來蒐集資料之品質：</p> <ol style="list-style-type: none"> 1. 問卷設計：如是否依據理論與分析的需求，事先擬定內容大綱？問題設計是否嚴謹？ 2. 調查訪問期間：建立各項執行過程的標準化作業流程。例如：訪員訓練、訪員手冊的編撰、問卷複查等。 3. 資料整理期間：建立資料檢誤的標準流程。例如：資料不合理值檢查、邏輯檢查、更正資料方式等的流程。 4. 相關文件紀錄：設計表單記錄執行過程，例如：訪視紀錄、問卷回收紀錄、複查紀錄、檢誤紀錄、資料修改紀錄等。
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	1.8	新蒐集的資料與現存資料的關係	確認現存資料是否可能有助於新資料的建立。
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	1.9	將如何整合現存資料與新蒐集的資料？	可能包括：與現存資料格式相容性的技術整合、資料信度及資料品質的議題等。
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	1.10	相較於現存資料，新蒐集的資料價值為何？	可能的價值包括：細節更多、範圍更大、驗證現存資料等。
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	1.11	後設資料（Metadata, 包括資料產生的過程細節，資料的架構，及資料的內容說明）的建立規劃	<p>國際間專為社會科學調查資料設計的后設資料標準格式為 DDI (Data Documentation Initiative)，可分為五部份：</p> <ol style="list-style-type: none"> a. 文件描述：提供撰寫引用來源 (citation) 的資訊，包括計畫名稱、執行機構、研究經費來源、資料提供單位、資料版本，以及本計畫參考的文獻資訊等

			<ul style="list-style-type: none"> b. 研究描述：計畫名稱、摘要、關鍵詞、執行機構、抽樣方法、抽樣誤差、調查期間與方法、調查範圍、研究方法、完訪率、資料存取與釋出條件等 c. 資料檔描述：資料檔檔名、資料格式、變項數、觀察值總數、遺漏值、權數等 d. 變項描述：問項、答項、變項說明、選項數值說明、有效數值範圍、次數分配、變項主題詞等 e. 其他相關文件：例如訪員手冊、協助受訪者回答問題使用的卡片、抽樣程式、相關技術報告等 f. <u>參考國外網站 (DDI)</u>
	2	研究倫理與隱私議題	
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	2.1	<p>蒐集的資料會因為牽涉到受訪者隱私或研究倫理，以致資料無法提供他人使用，與人共享嗎？</p>	<p>若資料蒐集可能造成受訪者的風險，可能需要移送部門或機構的研究倫理委員會 (IRB) 討論，並取得正式的受訪者同意書。</p> <p>如果欲蒐集的資料涉及受訪者隱私問題，可能的解決方法包括：</p> <ul style="list-style-type: none"> a. 資料匿名：在研究期間或在準備釋出與人共享的過程中，可刪除個人隱私資料或可識別出受訪者身份的資訊。 b. 資料分級：如可依資料敏感性高低分為公開使用版、限制使用版及現場使用版(on-site)，並分別訂定使用資格、取得及使用資料方式。 c. 重新編碼(recoding)或將資料重建(reconstruction)：將次數較少的類別合併到其他類別中，或將連續變項在一定數值以上或以下的極端

			數值，改以較不極端的數值取代。
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	2.2	<p>是否取得受訪者知情同意書 (informed consent)？提供的資訊是否足夠讓受訪者充份瞭解？同意書的內容是否包含資料共享？以及如何進行資料共享？</p>	<p>受訪者簽署的同意書內容和個資法的規定會影響資料儲存方式、有權查閱/使用的人員身份、以及可保存時間等。知情同意書可分為「研究說明」(information sheet) 及需要受訪者簽名的「同意書」(consent form) 兩部份，應在「研究說明」中提供充份的背景資料及資料共享資訊，供受訪者瞭解：</p> <p>a. 背景資料應包括：計畫基本資料、受訪者參與的研究項目、預期效益及可能風險、退出權利及退出時已提供資料的處理方式、資料使用方式(研究期間的使用分析；計畫結束後的儲存和資料共享)、確保資料使用符合研究倫理的措施(維護個人資料的機密性、匿名處理)。</p> <p>b. 資料共享的資訊包括：資料是否提供其他研究者使用？提供方式為何？是否交由專業資料庫典藏？如何確保共享時的資料被合理使用(是否需要限制使用者身分、限制的目的)、資料庫保護資料安全的措施。</p> <p>c. <u>參考網站 1</u>、<u>參考網站 2</u></p>
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	2.3	<p>資料取得是否符合個資法及其他相關法令規定？若有，採取哪些措施以符合法令規定？</p>	<p>建議可諮詢法律專家。</p> <p>2.1 有提供一些保護受訪者隱私資料的方法。</p>
	3	資料分享	
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	3.1	<p>資料是否會公開提供他人分享？何時開始提供分享？</p>	<p>若有任何資料分享上的限制，都必須是合理的。若計劃分享，應擬定具體的釋出時間表，以儘可能維護經費補</p>

			助單位肩負的公共利益責任。 影響資料能否公開及公開時間早晚的因素，可能包括：因申請專利而延後公佈資料的時間；為使計畫主持人保有優先利用資料發表著作的權利，資料可先保存一段時間後再公開；依據商業合作協議而無法將資料與他人分享。
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	3.2	若不將資料提供分享，原因是？	不分享的原因，可能包括：道德倫理因素、與資料提供者的保密協定、品質相關議題。(如果只提供部分資料，說明原因及可提供的資料)。
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	3.3	若可將資料提供分享，請說明分享的管道	說明如何將資料共享，例如：將它們存放在某個資料庫(中心)、有興趣的人可向你索取轉拷。 但後者需設法讓別人知道有這份資料存在，可能的方法包括：透過研究成果的出版、在學術會議上發表論文、將研究資料發表在網站上。
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	3.4	若可將資料提供分享，其他研究者何時可以開始使用資料？	a. 資料蒐集/建置/研究者是否有優先使用期？如果是，優先使用期間有多長？ b. 是否有任何政治、經濟、或專利考量，而影響資料的開放使用時間？
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	3.5	若可將資料提供分享，需要收費嗎？如要收費，說明收費細節	
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	3.6	哪些團體或組織會對此資料感興趣	
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	3.7	資料未來可如何被再利用	例如，可能用於哪些研究或議題。
	4	計畫執行期間的資料儲存及管理	
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	4.1	計畫執行期間，資料將儲存在哪裡？主要的儲存媒體	a. 考量儲存環境和儲存媒體 b. 考量儲存媒體的損壞風險

		為何？(包含數位化資料與非數位化資料)	<p>c. 使用的資料檔案格式是否長期有效？是否定期更新？</p> <p>d. 資料傳輸的方式，是否加密？</p>
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	4.2	是否建立電子檔案(包含文件電子檔、資料檔)？檔名的分類及命名規則？資料檔變項的命名規則？	<p>設計命名規則時可考量下列幾項原則：</p> <p>a. 可從檔案名稱分辨資料類別，例如是數據資料檔案或文件檔案；而文件檔案又可再進一步識別出是哪一類型的文件（例如：報告書？問卷？紀錄表單？）</p> <p>b. 可從檔案名稱分辨出檔案的格式，例如是採用哪一種軟體建立的</p> <p>c. 可將日期納入檔名，以分辨出版本新舊</p> <p>d. 若為長期追蹤調查或同時調查不同對象時，可設計能從檔名或變項名稱中分辨出追蹤波次或調查對象的規則</p> <p>e. 變項名稱能反映出相對應的問卷題目題號，或是題目內容</p> <p>f. 變項說明與選項數值說明的縮簡方式需有固定規則</p>
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	4.3	當轉換資料格式時，是否檢查資料完整性？資料檔內的後設資料是否遺失或被更改？	<p>需檢查的包括：</p> <p>a. 轉換成不同統計軟體格式時，資料是否遺漏、出現亂碼</p> <p>b. 將完訪的紙本問卷 keyin 成資料檔時，資料是否遺漏</p>
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	4.4	如何掌握資料版本資訊？	<p>可參考以下方式：</p> <p>a. 可指派專人或由文件管理系統統一管理，或將版本資訊放入檔案名稱中以利辨識。</p> <p>b. 利用主檔(master file)進行管理：「主檔」是用來儲存最完整與最新資訊的主要資料檔，需要永久保</p>

			<p>存。而為了因應不同目的之使用需求，每次可從主檔中摘錄、或選取部份資料，直接引用或做為範本加以修改，以符合當次使用之目的。</p> <p>c. 資料同時保存於不同媒體或場所時，應設計同步更新資料的機制，以確保版本一致性。</p>
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	4.5	在計畫執行期間，如何備份資料？	<p>a. 考量備份成本與儲存位置</p> <ul style="list-style-type: none"> ● 設備故障、遺失及損毀的風險(可採用 RAID1 或 RAID5 機制或是進行異地備援) ● 人工備份(專人負責)或利用軟體自動備份 <p>b. 備份頻率的規劃</p> <p>c. 備份資料應能確實還原</p> <p>d. 若有多個資料備份，復原時的優先順序</p>
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	4.6	如何實施資料權限控管及確保資料安全？	<p>a. 是否制定詳細的規則來管理資料？</p> <p>b. 數位、非數位資料和任何備份資料，保存在安全地點。</p> <p>c. 透過各級密碼管理來保護資料安全。</p> <p>d. 誰可以存取哪些資料，是否有相關的規定？</p>
	5	計畫結束後的資料儲藏、保存及管理	
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	5.1	哪些資料會被長期保存？	<p>a. 資料保存機構是否有收錄的規定或標準？</p> <p>b. 資料應全部保存、部分保存、或是完全不保存資料。</p> <p>c. 如何篩選保存的資料種類。</p> <p>d. 是否保留原始資料、抽樣、加權或其他計畫等相關資料。</p>
<input type="checkbox"/> 已具備	5.2	如何長期管理資料？	建議交由專業機構長期保存。

□N/A			如果資料包含限制性或敏感性資料，則應主動向管理機構提出限制使用資格及使用方式，例如：需要額外申請使用資料、或只能在指定地點使用等
□已具備 □N/A	5.3	資料是否進行隱私資料刪除、匿名處理或是重新編碼？	
□已具備 □N/A	5.4	是否將資料提供給專業機構進行長期保存？	<ul style="list-style-type: none"> a. 經費補助單位可能有選定的資料保存地點。 b. 是否與資料保存機構協商，確保資料被適當的保存？ c. 是否授權資料保存機構決定資料釋出的方式及規格？ d. 是否授權保存機構，由後者全權決定機敏資料（如受訪者聯絡資訊）能否提供其他研究者在合法的情況下再次使用？
□已具備 □N/A	5.5	若資料為自行保管，資料保存多久？	經費補助單位或所屬研究機構可能有指定的資料保存期限。
□已具備 □N/A	5.6	如果自行保管資料，長期保存的規劃為何？	<ul style="list-style-type: none"> a. 如果自行保管資料，應制定長期管理計畫，例如是否將資料釋出、如何長期保存以及怎麼管理。 b. 誰可以存取哪些資料，是否有相關的規定？
□已具備 □N/A	5.7	為了增加將來資料再度被使用的機會及使用時的便利性，應事先考量需要準備的後設資料及參考文件，並與資料一起保存	<ul style="list-style-type: none"> a. 若是交由專業機構保存，應瞭解該機構訂定的後設資料標準，並依標準製作後設資料。 b. 參考文件包括：檔案格式說明、過錄編碼簿、使用者手冊、研究報告、參考文獻、研究計畫書等。
	6	管理需求與確認	
□已具備 □N/A	6.1	執行資料管理計畫的人力安排，以及各自的角色及責任	可能包括資料管理時間的分配、技術方面的專案管理、人員培訓需求、儲存與備份及其他相關事項。

<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	6.2	資料管理工作所需的人力及經費預估	
<input type="checkbox"/> 已具備 <input type="checkbox"/> N/A	6.3	如何確認有無遵循資料管理計畫？是否訂立查核檢討的時間表？	